A support vector machine for predicting spontaneous termination of paroxysmal atrial fibrillation episodes

A Support Vector Machine for Predicting Spontaneous Termination of Paroxysmal Atrial Fibrillation Episodes

JD Diaz[1,2], C Gonzalez[2], O Escalona[2]

[1]Francisco de Miranda University, Coro, Venezuela
[2]Simon Bolivar University, Caracas, Venezuela

Abstract

The aim of this work is to predict the spontaneous termination of atrial fibrillation (AF) episodes. The database includes three record groups: non-terminating AF (N), AF that terminates one minute after recording end (S), and AF that terminates immediately after recording end (T). A first goal consisted on separating N from T group records (event 1), and a second, for separating S from T records (event 2). A Support Vector Machine was used for the classification problem. For event 1, four indexes were extracted: the atrial fibrillatory frequency (AFF) and the mean, standard deviation, and approximate entropy of RR intervals. For event 2, the AFF, the energy of the 3-7 Hz and 7-11 Hz bands, from the ten and five final seconds of the records, were used. The groups were divided in two sets: learning and test. For event 1, a 100% in learning, and 86.66% in test set were correctly classified. For the event 2, we classified 100% in the learning, and 80% in the test set.

1. Introduction

Atrial Fibrillation (AF) is the most common cardiac arrhythmia, affecting 1% of the population [1]. AF is characterized by predominantly uncoordinated atrial activation with consequent deterioration of atrial mechanical function. On the electrocardiogram (ECG), it is indicated by irregular fibrillatory waves, giving rise to a loss in the normal P wave, which represents the atrial activity.

When atrial fibrillation appears in its paroxysmal (self-terminated) form (PAF), it may recur with a variable frequency over many times. In many patients the arrhythmia may undergo transition to the persistent (chronic, non-self-terminated) form (CAF). The transition rate varies depending on the etiology. The duration of the paroxysmal events also influences over the transition rate [2].

Changes in action potential duration and atrial refractory periods, are indicators of atrial electrical remodeling after an atrial fibrillatory event. This atrial remodeling may play an important role in the self-perpetuation of AF [3, 4].

CAF requires pharmacological treatment or electric shock delivery, in order to be finished. Sustained AF has been associated as a factor that increases the risk of stroke and thrombus genesis [5, 6].

A better understanding about the mechanisms of self-termination of atrial fibrillation and its perpetuation, may contribute to choose an adequate therapy, leading to decrease risks in patients and costs of AF treatment.

Differences between PAF and CAF on superficial electrocardiogram are not evident, therefore, predicting the end of spontaneous atrial fibrillation using ECG represents a challenge. In 2004, Physionet and Computers in Cardiology asked if it is possible to predict if (or when) an episode of AF will end spontaneously [7].

The aim of this work was trying to predict the spontaneous termination of AF episodes, as it was raised in the Physionet/Computers in Cardiology Challenge 2004.

2. Methods

2.1. Database

The database included three record groups of electrocardiograms from the 2004 PhysioNet/Computers in Cardiology Challenge database [8]:

Group N: Non-terminating AF, at least an hour following at the end of the recording,

Group S: AF that terminates one minute after recording end, and

Group T: AF that terminates immediately (within one second) after recording end.

The database is composed of 80 one minute surface ECG records. That data was extracted by a two-channel Holter ECG with sampling rate of 128 Hz and 16 bits resolution. DII and V1 patient leads were recorded.

The database was divided into a learning set and two test sets (test set A and test set B). The learning set contains 30 records in all, with 10 records in each of three groups. Test set A contains 30 records of which about half are from Group N and the remainders are from Group

T. Test set B contains 20 records, 10 from Group S and 10 from Group T.

2.2. ECG signal processing

Two goals were raised. The first of them consisted on separating records of group N from group T (event 1), and the second of them, was carried out for separating records of group T from group S (event 2). All signal processing was done in Matlab (The MathWorks, Inc., Natick, MA, USA).

Event 1: We applied an approach for computing four characteristic indexes from the ECG signal on this event: the atrial fibrillatory frequency (AFF), and from the RR intervals time series, the RR mean value (RRmean), standard deviation (RRsd) and approximate entropy (ApEn).

The dominant atrial fibrillatory frequency, estimated using power spectral analysis on the atrial fibrillatory activity in the ECG, reflects the average rate of AF [9, 10]. This index is useful for non-invasive assessment of electrical remodeling in AF [11]. Several investigations have shown that longer paroxysmal AF episodes have higher frequencies than shorter ones, and patients with low AFF are more probably to be cardioverted using antiarrhythmic drugs [9, 12, 13].

Frequency analysis was carried out through three steps: band-pass filtering, ventricular activity cancellation (QRST complexes), and fast Fourier transformation (FFT). Band-pass filtering of the data was performed using a 0.5-50 Hz, high order, zero-phase filter. This was chosen to reduce both respiration induced fluctuation of the baseline and electrical noise. The original signal was up-sampled to a 1024 samples per second in order to have more signal definition. QRST complexes were subtracted using a template matching and signal averaging. QRST segments were classified according to their morphology, and then averaged. A template was created with the QRST segments averaged, and the transitions between successive QRST complexes in the reference signal were the means of corresponding intervals of the original signal. The template was subtracted from the original signal to produce a residual atrial fibrillatory signal [9, 10, 14].

Residual atrial fibrillatory signal was down-sampled to 256 Hz. The power spectrum of the residual signal was calculated by 4096 points windowed FFT, 1024 points Gaussian window and 853 points overlap. An average was made over the 60 seconds of the power spectrum of the residual signal. Peak frequency was determined in the 3-12 Hz range, and the AFF was estimated as the component with maximum amplitude. This last step was only made over the signal of lead V1. Lead V1 was chosen because it had the largest amplitude of fibrillatory activity according to the results of previous studies [9, 10].

Figure 1, illustrates the different steps described for getting the residual atrial fibrillation activity for 10 seconds. The template (green signal in the middle), is subtracted from the filtrated original signal (the blue one on top) to produce a signal with ventricular activity suppressed (the red signal on bottom).

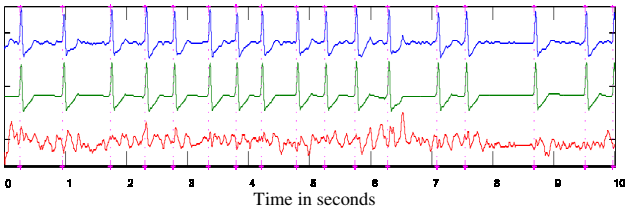

Figure 1. Ten seconds from signal filtered (blue), template (green), and atrial fibrillation residual signal (red).

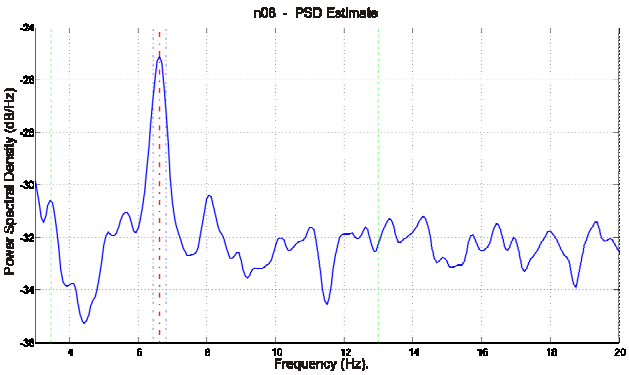

Figure 2. Power spectral density of the 60 sec. of the atrial fibrillatory activity signal. Peak represents the atrial fibrillatory frequency.

In Figure 2, we can observe the power spectrum average of the 60 seconds of the residual signal.

The mean and standard deviation of the interbeat period series were computed for the entire ECG signal. The approximate entropy of the RR intervals was also calculated.

Approximate entropy corresponds to a measurement of complexity and disorder of a data series or a signal. This was calculated using the Pyncus method [15]. Two parameters must be specified before ApEn can be computed: m, the embedding dimension of the vectors to be formed, and r, a threshold that is, in effect, a noise filter. As suggested by Pincus, we used m=2 and r=0.25SDx, where SDx is the standard deviation of the original data.

Event 2: On this approach, our hypothesis was that the 10 and 5 final seconds of the recording, could show

significant differences for the AFF and the percentages of energy of the 3-7 Hz and 7-11 Hz bands of the residual atrial fibrillatory signal, between the group which AF terminates immediately (group T), and the group that AF terminates one minute after recording end (group S).

FFA was computed like in event 1, but the power spectrum was only calculated over the 10 final seconds of the residual atrial fibrillatory signal.

The percentages of energy of the 3-7 Hz and 7-11 Hz bands were calculated using *wenergy* function of Matlab, which computes the percentage of energy corresponding to the approximation and details of a wavelet decomposition. A *Symlet*, fourth order wavelet filter was used for the wavelet decomposition. The energy percentages were computed over the ten and five final seconds of the residual atrial fibrillatory signal for each record in this event.

2.3. Statistical analysis

All continuous variables are presented as mean ± 1 standard deviation (SD). These indexes were compared between groups for each event. Nonparametric Wilcoxon unpaired test is applied to analyze differences between the record groups.

2.4. Records classification

A two order, polynomial function, support vector machine (SVM) was employed for the classification problem. For each event, a SVM was only trained using the indexes computed from learning sets of the corresponding event.

The parameters of the test sets were used for evaluating the performance of each SVM trained.

The SVMs were implemented using libraries designed by the Taiwan University [16]. A Kernel of polynomial function was employed for the SVM implementation, using the equation 1:

$$(\varphi \cdot V' \cdot V + r)^d \quad (1)$$

where V represents the vector of variables, d sets degree of the kernel function, and φ and r are coefficients for setting the SVM. We adjusted d=2, φ=2 and r=1 in the SVMs built.

3. Results

3.1. Event 1:

Statistical analysis revealed significant differences between group N and group T for the parameters FFA and ApEn of RR intervals. There were not significant differences in the RRmean and RRsd between these groups.

Tables 1 and 2, show the statistical analysis results, obtained on learning and test sets, for each computed parameter for the record groups in this event.

Parameter	Group N mean±SD	Group T mean±SD	P-value
AFF (Hz)	6,58±0,62	5,11±0,35	NS
RRmean (sec)	0,80±0,17	0,64±0,21	0,112
RRsd (sec)	0,17±0,06	0,16±0,07	0,41
ApEn	0,56±0,10	0,65±0,13	0,048

Table 1. Statistical analysis of the computed parameters in learning set. NS- non significant

Parameter	Group N mean±SD	Group T mean±SD	P-value
AFF (Hz)	6,57±0,78	5,21±1,16	0,001
RRmean (sec)	0,71±0,12	0,65±0,18	0,294
RRsd (sec)	0,15±0,06	0,14±0,06	0,589
ApEn	0,54±0,11	0,67±0,18	0,025

Table 2. Statistical analysis of the computed parameters in test set

We used different combinations of the parameters for training the SVM, getting different performances in each case. Table 3 resumes the gotten results for the classification.

Parameter employed	Exactitude learning set	Exactitude testing set	Total exactitude
AFF	95% (19/20)	80% (24/30)	86% (43/50)
AFF and ApEn	100% (20/20)	80% (24/30)	88% (44/50)
AFF, ApEn, RRmean, and RRsd	100% (20/20)	86,66% (26/30)	92% (46/50)

Table 3. Classification exactitude in learning and test sets, for different combinations of the parameters used for training the SVM

3.2. Event 2:

Statistical analysis did not reveal significant differences in none of the calculated parameters between the groups S and T, in both learning and test sets.

Nevertheless, we could get a total exactitude of 90% in the classification of the groups on this event. This best score was obtained by training the SVM with the following parameters: AFF computed from the 10 final seconds of the recording, the percentages of energy of the

3-7 Hz and 7-11 Hz bands from the 10 final seconds, and the percentage of energy of the 7-11 Hz band of the 5 last seconds.

The SVM correctly classified 20/20 records (100%) in learning set, and 16/20 records (80%) in test set.

4. Discussion and conclusions

More organized atrial fibrillation, showing single wavefronts propagating through atrial myocardium, has a longer atrial fibrillatory cycle average than less organized AF with multiple wavefronts and areas of conduction block [17]. Then, the average of atrial fibrillatory cycle length (inverse of atrial fibrillatory frequency) is an index of atrial refractoriness and atrial fibrillation organization. On the other hand, longer AF episodes reverting spontaneously, have a higher fibrillation frequency value than shorter episodes [9, 18]. Thus, higher AFF suggests more disorganized atrial fibrillation, and it has less probability of finishing spontaneously in short time.

Approximate Entropy can be used by predicting the end of atrial fibrillation episodes. In this study we could find that the RR intervals show a higher disorder when AF is near to its end than when it is far of it.

It is possible to predict the spontaneous termination of atrial fibrillation by employing artificial intelligent techniques, like support vector machine, but evidently, it must use parameters for training, which are able to characterize the system that is wanted to predict, in order to get the best performance.

Acknowledgements

This work was partially supported by Francisco de Miranda University and Simón Bolívar University, Venezuela.

References

[1] Waktare J. Atrial Fibrillation. Circulation 2002. 106:14-16.

[2] Nunain S, Debbas N, Camm A. Determinants of the Course and Prognosis of Atrial Fibrillation, pp: (28) 350-358, in: Touboul P, Waldo AL (eds). Atrial Arrhythmias: Current Concepts and Management. Mosby Year Book Inc, 1990; 1

[3] Daoud E, Bogun F, Goyal R, Harvey M. et al. Effect of Atrial Fibrillation on Atrial Refractoriness in Humans. Circulation. 1996. 94:1600-1606

[4] Wijffels M, Kirchhof Ch, Dorland R, Allessie M. Atrial Fibrillation Begets Atrial Fibrillation: A Study in Awake Chronically Instrumented Goats. Circulation. 1995. 92:1954-1968

[5] Wolf P, Abbott R, Kannel W. Atrial fibrillation: a major contributor to stroke in the elderly. The Framingham Study Arch Intern Med., 1987. 147:1561–1564.

[6] Halperin JL, Hart RG. Atrial fibrillation and stroke: new ideas, persisting dilemmas. Stroke 1988;19:937– 41.

[7] Moody G. Spontaneous Termination of Atrial Fibrillation: A Challenge from PhysioNet and Computers in Cardiology 2004. Computers in Cardiology, 2004: 101-104.

[8] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM,Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK,Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000: 101(23):e215-e220

[9] Bollmann A, Kanuru N, McTeague K, Walter P, DeLurgio D, Langberg J. Frequency Analysis of Human Atrial Fibrillation Using the Surface Electrocardiogram and Its Response to Ibutilide. Am. J. Cardiol. 1998. 81:1439–1445

[10] Holm M, Pehrson S, Ingemansson M, Sornmo L, Johansson R, Sandhall L, Sunemark M, Smideberg B, Olsson C, Olsson S. Non-invasive assessment of the atrial cycle length during atrial fibrillation in man: introducing, validating and illustrating a new ECG method. Cardiovasc. Res. 1998. 38:69-81

[11] Husser D, Stridh M, Sornmo L, Olsson B, Bollmann A. Frequency Analysis of Atrial Fibrillation From the Surface Electrocardiogram. Indian Pacing and Electrophysiology Journal 2004. 4(3):122-136

[12] Fujiki A, Nagasawa H, Sakabe M, Sakurai K, Nishida K, Mizumaki K, Inoue H. Spectral characteristics of human atrial fibrillation waves of the right atrial free wall with respect to the duration of atrial fibrillation and effect of class I antiarrhythmic drugs. Jpn Circ J. 2001. 65:1047-51

[13] Biffi M, Boriani G, Bronzetti G, Capucci A, Branzi A, Magnani B. Electrophysiological effects of flecainide and propafenone on atrial fibrillation cycle and relation with arrhythmia termination. Heart 1999. 82:176–182

[14] Bollmann A, Mende M, Neugebauer A, Pfeiffer D. Atrial Fibrillatory Frequency Predicts Atrial Defibrillation Threshold And Early Arrhythmia Recurrence In Patients Undergoing Internal Cardioversion Of Persistent Atrial Fibrillation. Journal Of Pacing And Clinical Electrophysiology 2002. Vol. 25, No. 8:1179-1184

[15] Akay M. Nonlinear Biomedical Signal Processing, Dynamic Analysis and Modelyng. IEEE Press Series on Biomedical Engineering 2000. Vol. 2.

[16] Chih-Chung Ch, Chih-Jen L. LIBSVM: a library for support vector machines. 2001. Software vailable: http://www.csie.ntu.edu.tw/~cjlin/libsvm

[17] Konings K, Kirchhof C, Smeets J, Wellens H, Penn O, Allessie M. High-density mapping of electrically induced atrial fibrillation in humans. Circulation 1994 89:1665-80

[18] Bollmann A, Sonne K, Esperer H, Toepffer I, Langberg J, Klein H. Non-invasive assessment of fibrillatory activity in patients with paroxysmal and persistent atrial fibrillation using the Holter ECG. Cardiovasc Res. 1999. 44:60-6

Address for correspondence

José D. Díaz R.
Grupo de Bioingeniería y Biofísica Aplicada.
Universidad Simón Bolívar.
Valle de Sartenejas, 89000.
Caracas, Venezuela
E-mail: 03-83260@usb.ve